Authoritative Data and Crowd Sourcing: What's the Connection?

or

How Can Volunteers Provide Authoritative Data to the BGN?

David Stage

Authoritative Data and BGN Business Requirements

Authoritative source:

- An organization that has legislative mandate and "budget" to create data to meet a specific business need

- **Place names for USGS quads**

- **Provide an official process to create and maintain official place names**

- **Maintain currency of that data**

- **Maintain place name history**

The National Map Corps
A door for "crowd-sourced" data to become authoritative data

Authoritative Data Sphere
Inside the Process

- **Authority:** Legal responsibility by a public agency to conduct business for the public good
- **Authoritative data source:** has legal authority to collect data <u>for a specific business purpose</u>
- **Authoritative data** is officially recognized data that comes from an authoritative source
- **Data steward**: entity within source that has the responsibility for collecting data
- **Certified data:** has been vetted by an a authoritative entity and has legal standing
- **Uncertified data:** – more current but yet to be vetted

Volunteer Data Sphere
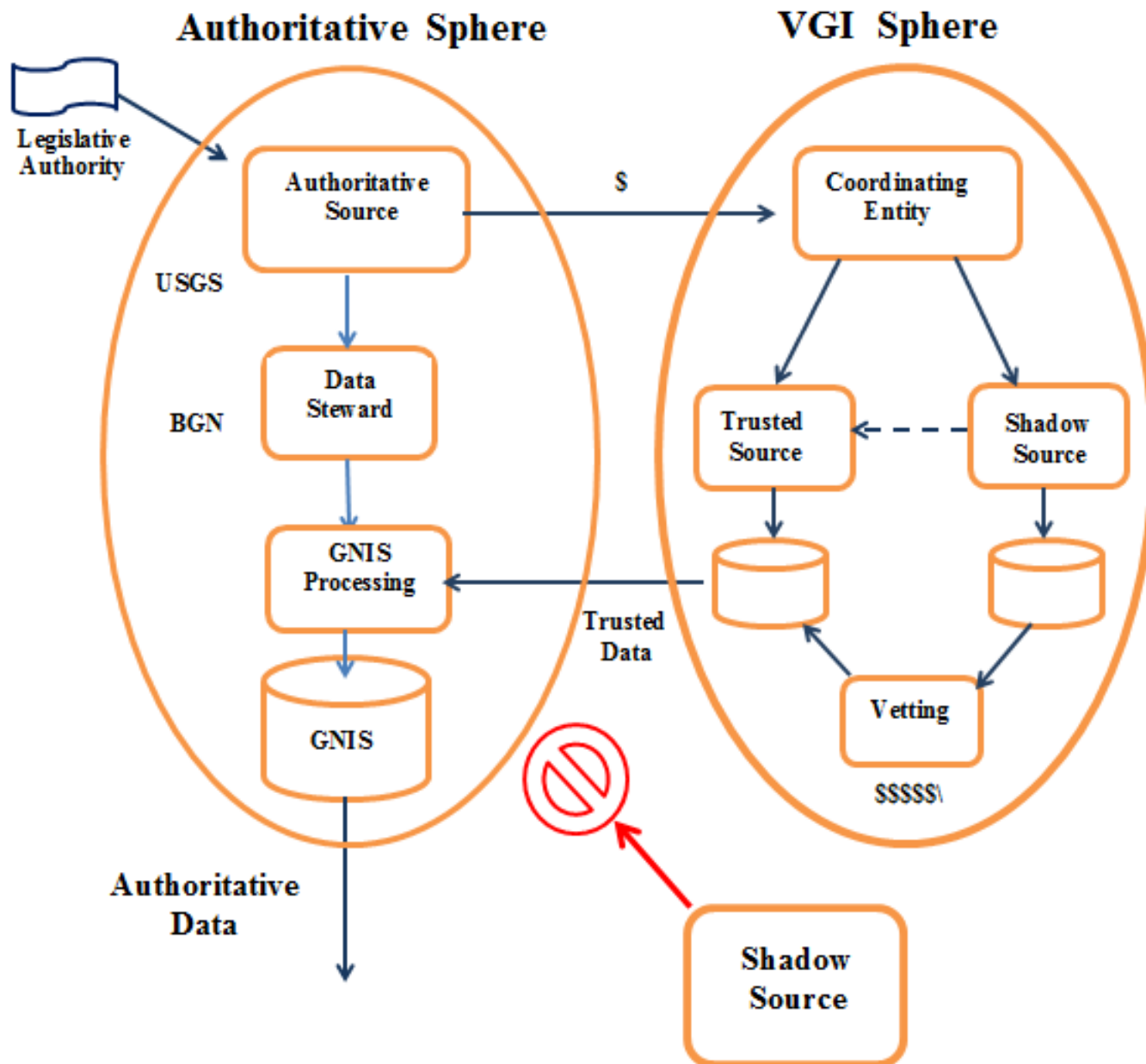## Sources outside of the process

- **Strangers of the Night**
  - **Shadow source:** Source whose veracity is unknown
  - **Shadow data:** Data from a source that is outside the authoritative process of which little is known
- **Friendships of Utility**
  - **Trusted source:** An entity whose trust has been earned
  - **Trusted data:** Data from outside the authorized process whose limitations, currency and attributes are known and can be verified

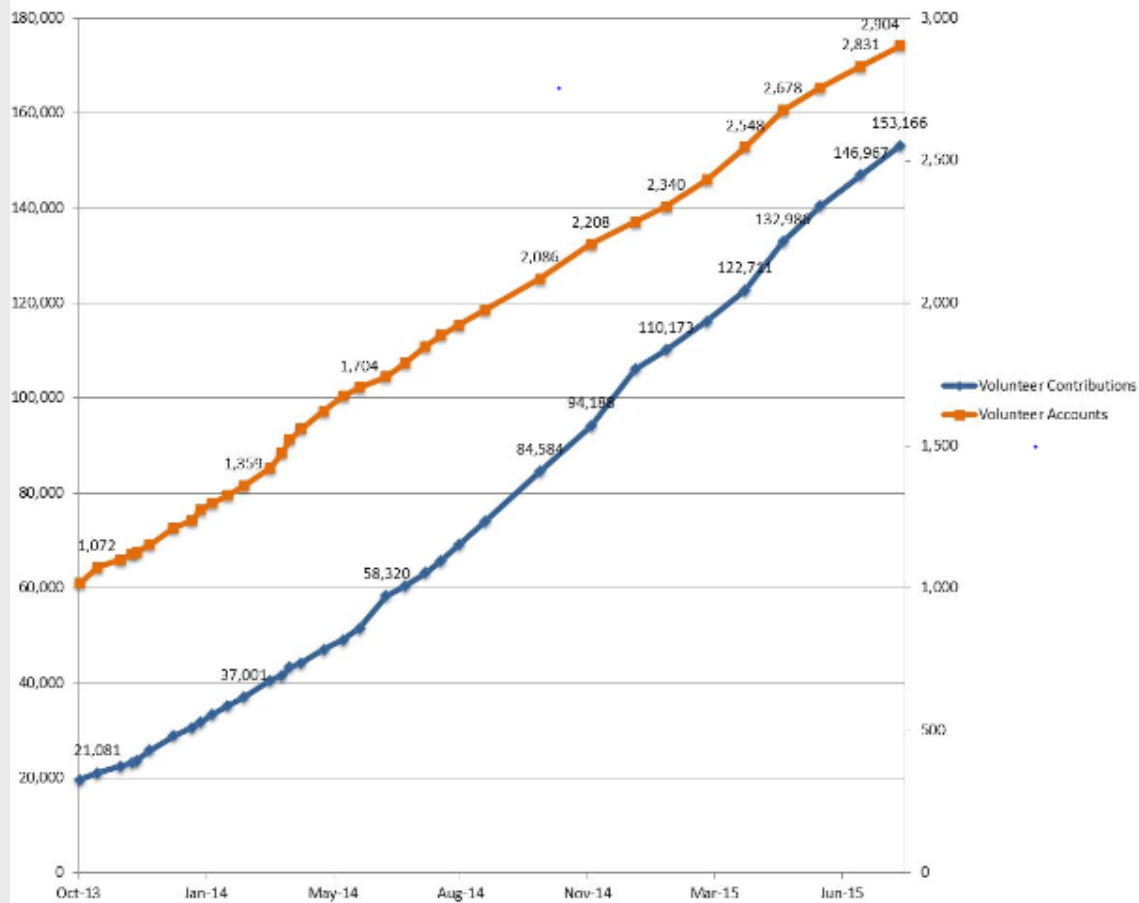*VGI – Volunteer(ed) Geographic Information

Volunteer Data Sphere
Volunteer Types

- **Type I:** Intergovernmental cooperatives
  - Direct management support
- **Type II:** NGOs, associations and government agencies
  - Indirect management support
- **Type III:** Independent individuals
  - Altruism
  - No infrastructure

Authoritative Sphere

VGI Sphere

Legislative Authority

Authoritative Source

$

Coordinating Entity

USGS

BGN

Data Steward

Trusted Source

Shadow Source

GNIS Processing

Trusted Data

GNIS

Vetting

$$$$$\

Authoritative Data

Shadow Source

The National Map Corps (TNMCorps)

2014 - 2015

- 125,000 Contributions
- 2,000 Accounts

The National Map Corps: Data Submitted by Provider

A universal phenomenon of volunteer efforts

Coordinators/Leaders Relationship to Volunteer

Goal: Create authoritative data using volunteers

What tools do you have?

- No authority over volunteers

- No funds

- No stick

- What can you do?

What to Do

A manager's job is to make it possible for their employees/volunteers to do a good job!

- **Motivation:** How can persons be motivated to volunteer their services?

- **Quality Assurance:** How do you facilitate quality control?

- **Trusted Sources**: How to transition *Strangers of the Night* to *Friends of Utility?*

Motivating Communities

Objective: Increase participation of volunteers

- **Provide recognition and rewards**

- **Make it easy** – minimal input and tools
  - **Training** (User Guides, WebEx, . . .)

- **Document success**
  - Pay the person to help you put together a training package

- **Target specialist communities** that have an interest in the resulting data (Type I – III Volunteers)

- **Accept the idea that going beyond your mission can be valuable motivator** (focus on volunteer community)

Quality Assurance
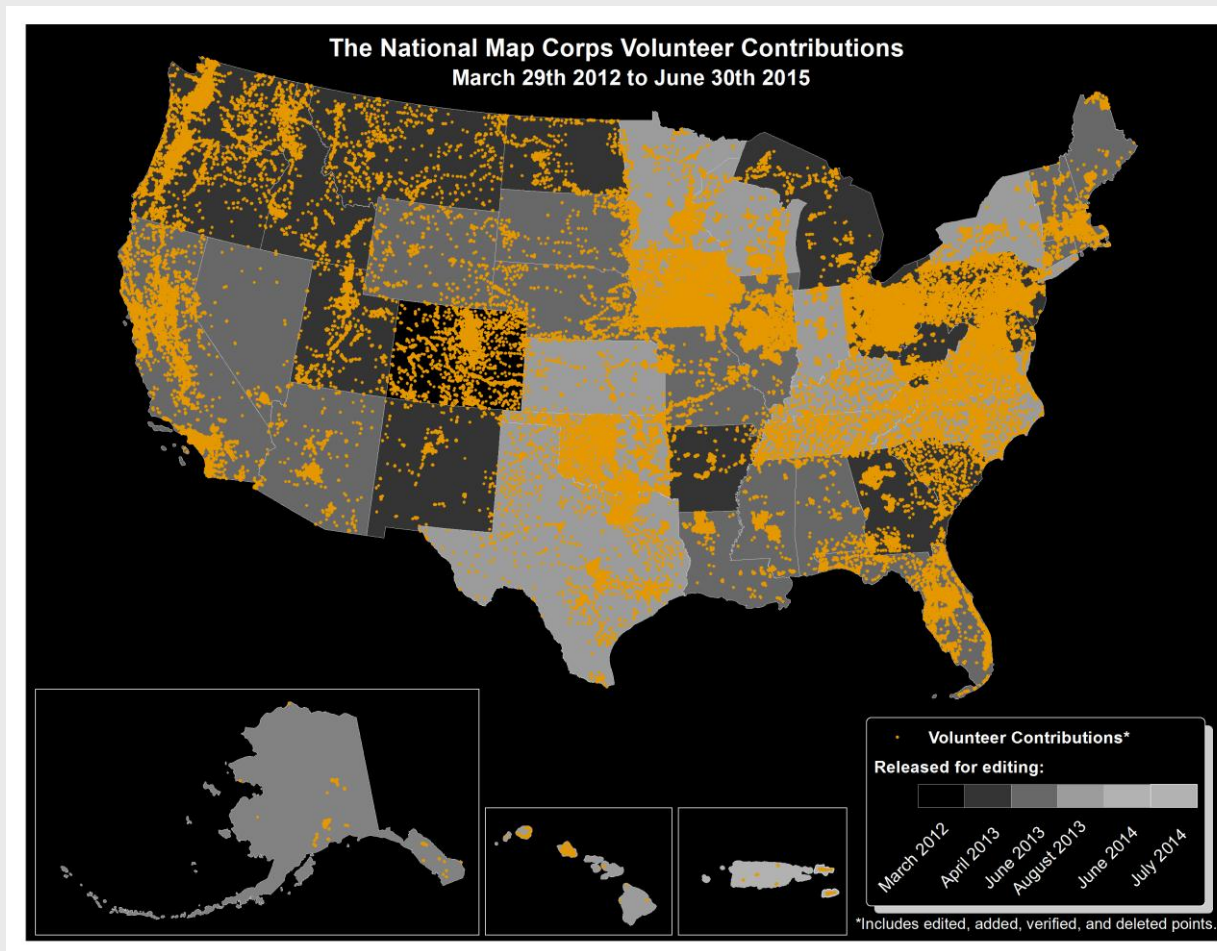## Misconception – Crowd-sourced data is unreliable

Objective: Assure data quality

1. **Self regulation (Wiki approach)**
   – Peer review
2. **Make it simple**
3. **Develop data checking procedures**
   – Sampling
   – Automation
4. **Track and evaluate data input by provider**
   – Identify quality data providers
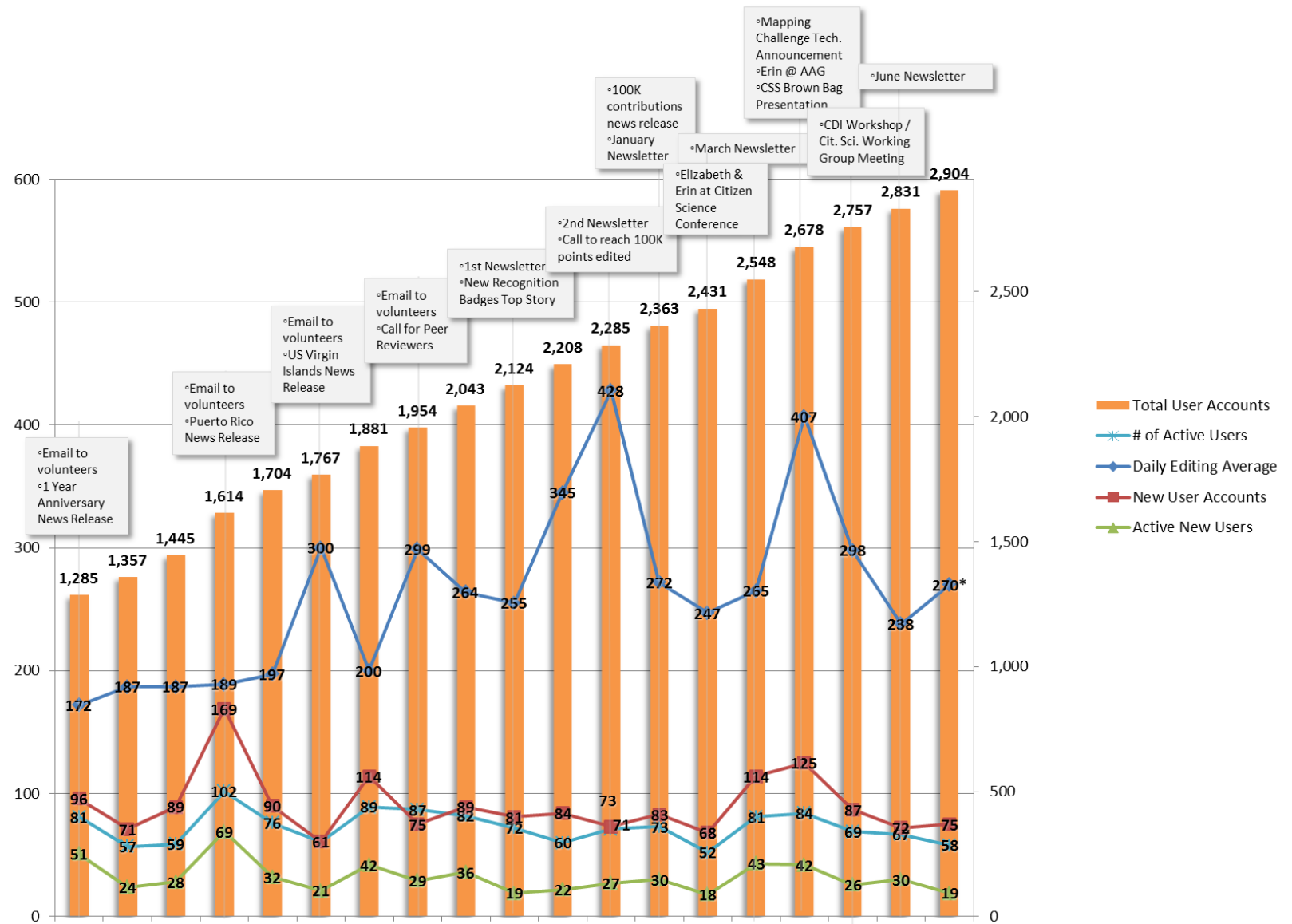   – Advanced Editor level is achieved by providing 200 good submissions

Creating Trusted Sources

Objective: increase confidence in data sources

- **Track and evaluate data input by provider**

- **Create levels of responsibility/authority based on the track record of the provider**

- **Provide super-provider with a fast-track pathway into the *Authoritative Data Sphere***
  - *TNMCorps Advanced Editor*

What's the Status?

The National Map Corps Volunteer Contributions
March 29th 2012 to June 30th 2015

Volunteer Contributions*
Released for editing:

March 2012 April 2013 June 2013 August 2013 June 2014 July 2014

*Includes edited, added, verified, and deleted points.

Lessons Learned

Laura Kostanski: *Crowd-Sourcing Geospatial Information for Government Gazetteers*

1. **Focus on the end-users**
2. **Don't rely solely on digital technology**
3. **Volunteers can only do so much**
   - Can't fill in all of the gaps
   - Be happy with what a user wants to provide
   - Provide guidance on scoped data sets and areas
   - Publish successful practices for targeted data
4. **Work with special interest groups (Type I, II and III)**
   - State and Federal agencies (emergency management, parks and recs, school systems, state GI boards)
   - NGO's (birders, hikers, canoeists, bikers, national, and state organizations)

Lessons Learned
Laura Kostanski

5. **Provide easy/simple accessible technology**
   - Don't make it complicated - ask for the minimum
6. **Develop policies for quality assurance**
7. **Track entries to identify trusted sources**
   - Allows provider to earn a higher trust status
8. **Provide greater responsibility to trusted sources**
9. **Give rewards/attention to volunteers**
   - Tangible: data, points, cash
   - Intangible: communication, improved information for user groups, more independence

Lessons Learned
Laura Kostanski

10. **Re-evaluate your data spectrum to attract user interest**

    – Provide unofficial data along with official data

        • Target data to user interest

        • Lines (trails), polygons/footprints

        • Publish successful procedures
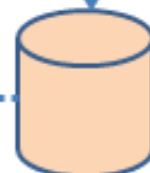
    – Offsite trusted sources

Authoritative Data Sphere

Trusted Data Sphere

BGN Guide Lines

Shapes

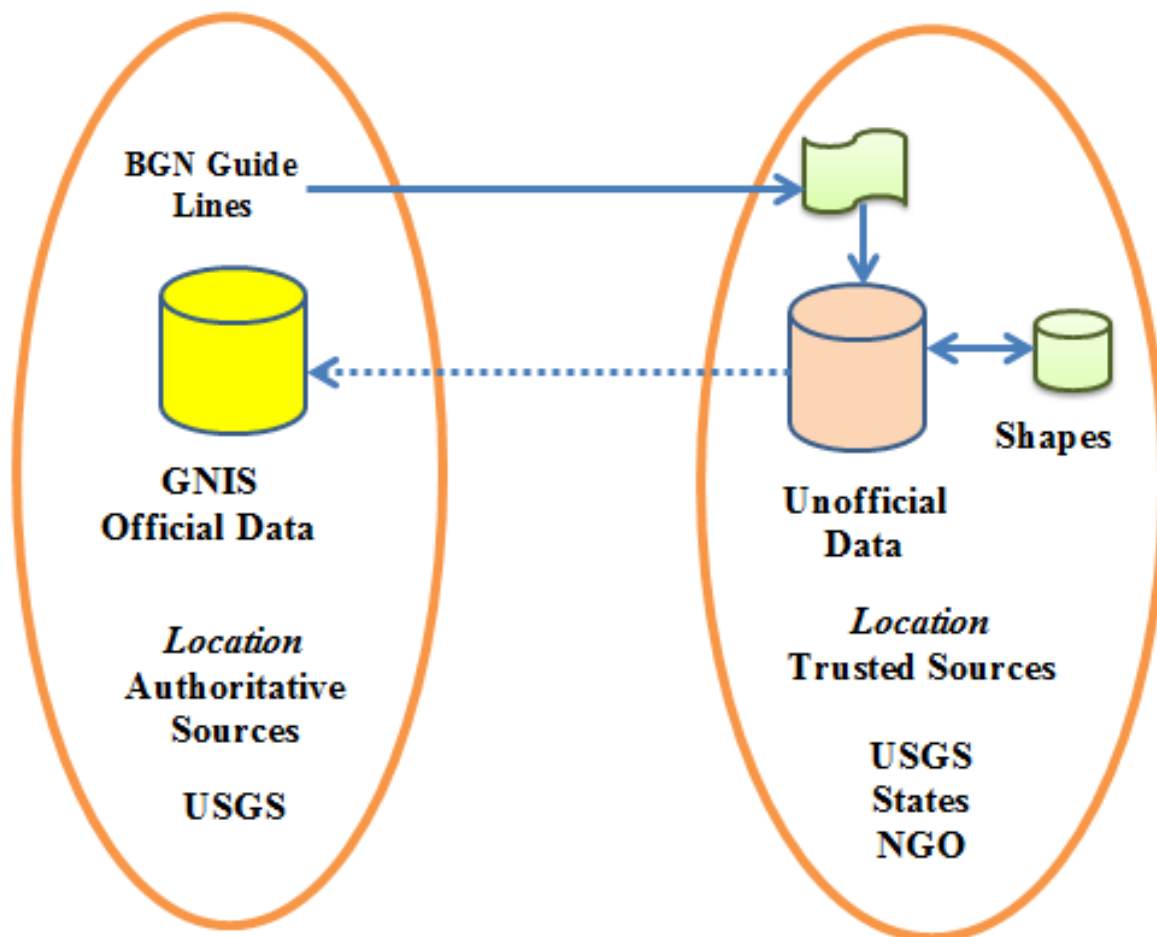GNIS Official Data

Unofficial Data

Location Authoritative Sources

USGS

Location Trusted Sources

USGS
States
NGO

Suggestions
("And Do That Too")

1. **User needs assessment** – organization by organization (LK 1, 4)*

2. **State GIS organizations** (GIS boards) and NGO's (NSGIC) (LK 1, 4)

3. **Challenges**: Identify tasks to be completed (LK 1, 5)

4. **Allow users to identify data of opportunity** (LK 8)

   – Provide high level volunteers with representative/liaison status

   *Laura Kostanski (LK) and number of lesson learned

Suggestions

5. **Provide access to unofficial and official data** (LK 10)

    – USGS, states, regional governments

    – Guidelines for unofficial data collections

6. **Footprints** (LK 10)

    – Lines & polygons

7. **Publish successful data collection procedures** (LK 5, 8, 9)

Conclusion

VGI can be used to create authoritative data

It's here now!

The future will be an interesting time!

Questions?

Key References

- Kostanski, Laura, Crowd-Sourcing Geospatial Information for Government Gazetteers, Submitted by Australia, Tenth United Nations Conference on the Standardization of Geographic Names, August 2012.*

- Elizabeth McCartney, Kari Craun, Erin Korris, David Brostuen and Laurence Moore; "Crowdsourcing the National Map", U.S. Geological Survey, National Geospatial Technical Operation Center, May 2015

- Erin Korris, Greg Matthews, Larry Moore, "Assessment of Data Quality Improvements in the National Map Corps Project for 2012 Colorado Pilot" (unpublished).

- David Stage, "Authority and Authoritative Data: A Clarification of Terms and Concepts", IAAO, Fair and Equitable, 2009, http://www.iaao.org/uploads/Stage.pdf